

Charles W. O'Donnell

Curriculum Vitae

Cambridge, Massachusetts
✉ charles.odonnell@gmail.com
<http://codonnell.org/>

Summary

Computational biologist with over ten years experience in multi-disciplinary, multi-institutional programs, including mentoring and direct management of scientists. Deep expertise in NGS assay design, analysis, and data integration, epigenomics, genetics, editing, gene therapy, regulatory cell circuits, modeling, systems and network analysis, public dataset integration, protein & RNA structure, high performance computing, and algorithms. Interest and experience ranging from early target discovery, mechanism-of-action studies, translation into human, and clinical studies.

Education

Ph.D., EECS, Massachusetts Institute of Technology (MIT), 2011.

Thesis: "Ensemble Modeling of β -sheet Proteins"

Thesis advisors: Bonnie Berger, Srinivas Devadas, & Susan Lindquist.

Minor in Biology.

S.M., EECS, Massachusetts Institute of Technology (MIT), 2005.

B.S., Computer Engineering, Columbia University, 2003.

Magna Cum Laude.

Research Experience

2014–present

Biogen, Senior Scientist, Research. *Epigenetics, Cell & Gene Therapy* (Nov. 2016 — present)

Biogen, Scientist II, Research. *Epigenetics, Cell & Gene Therapy* (Apr. 2014 — Nov. 2016)

- **Founding member of Biogen Epigenetics group (later merged with Cell & Gene Therapy).** Project and people manager, external partner liaison, NGS assay developer, and lead analytical contributor for Epigenetics projects — focusing on multi-omic data integration. Further investigated regulatory landscapes surrounding gene therapeutic targets of interest and completed internal Biogen management training.
- **Initiated and lead the establishment of Biogen Epigenome-wide association studies (EWAS).** Focused on the identification of monocyte-specific epigenomic features significantly associated with Parkinson's Disease, Alzheimer's Disease, and Trisomy 21. Worked closely across Research, Clinical, Digital Health (wearables), Business Development, and with hospital/clinic and academic partners resulting in active subject enrollment (public collab. Tel Aviv Sourasky Medical Center, Linda Crnic Institute for Down Syndrome). Team comprised of one computational biologist (direct report) and three bench scientists. Through these integrations of multidimensional NGS-based data, we aim to elucidate epigenomic modulation of GWAS-identified associations, identify markers of disease onset and progression, and discover novel, targetable epigenetic regulators upstream to pathophysiology.

- **Key individual championing R&D's development and expansion of NGS capabilities and multi-omics dataset integration.** Stepped forward to jointly formulate R&D's RNA-seq strategy, led build-out team, and implemented first centralized pipeline. Additionally introduced novel seq-based assays and other experimental technologies, developed methods for the analysis of epigenetic *-seq assays, piloted department-wide cloud computing (AWS) approaches, and contributed to internal standardization of multi-omic data integration and systems biological approaches.
- **Key member of Biogen/Rodin Therapeutics partnership exploring histone deacetylases (HDACs) as a therapeutic target in neurodegeneration.** Drove analytical assessments, NGS-related assay design, and contributed to strategic decision making throughout internal program development.
- **Investigating novel CRISPR-based therapeutic modalities for the modulation of epigenomic state at specific genomic loci/targets.** Analytical lead and initial project co-designer.

2011–2014

Harvard University, Postdoctoral Fellow, *Douglas A. Melton Lab (HSCI)*

MIT, Postdoctoral Fellow, *David K. Gifford Lab (CSAIL)*

NIH NRSA fellowship recipient

- **Modeling NGS data to uncover determinants of cell state and regulation.** I design Next-Gen-Sequencing experiments and develop new analytical tools to discover genetic and epigenetic factors that control cellular competence, determination, and differentiation. Working closely with other biologists in my labs, data from DNase-seq, RNA-seq, ChIP-seq, ChIA-PET, Bis-seq, and massively parallel reporter assays has been modeled to suggest effective stem cell differentiation and reprogramming strategies for the production of pancreatic β -cells and motor neurons (collab. with Wichterle Lab, Columbia University). I also manage our group involvement within the ENCODE project to apply these analyses to other systems and to work collaboratively across institutions on integrative methods for analyzing disparate ENCODE data sources.
- **Techniques for discovery of TF binding sites and chromatin architecture change.** A new algorithm was developed that predicts the genomic binding occupancy of hundreds of transcription factors (TFs) from a single DNase-seq assay. Implementation was done on Amazon's EC2 to take advantage of massively parallel efficiencies. We applied this method to discovery changes in chromatin state and TF binding across directed differentiation lineages in human and mouse ES cell lines within my labs and in collaboration with other labs (Wichterle Lab, Columbia University, Zhou Lab, Cowan Lab, Harvard University, Mannstadt Lab, HMS). In doing so a new class of directional pioneer TFs were discovered that can induce asymmetric chromatin accessibility change around a binding location. To extend this study a novel assay was developed that provides high DNase-seq coverage over select loci using sequence-capture and additional computational algorithms, providing a detailed look at chromatin structure at enhancer sites.
- **Regulatory interpretation of TF binding patterns.** The consequences of enhancer TF binding interactions (discovered via DNase-seq or ChIP-seq) are being studied using a join model that integrates RNA-seq expression data as well as 3-dimensional chromatin data such as ChIA-PET or Hi-C, which can link many enhancer regions to their target promoters after computational analysis. Binding patterns have been studied in human and mouse ES lines (in my labs and in collab. with Zhou Lab, Arlotta Lab, Harvard University), as well as a hematopoietic differentiation system (in collab. with the Zon Lab, HMS). Massively parallel reporter assays (MPRA) are being pursued to test the fidelity of enhancer structures in an integrated, iterative computational/experimental pipeline (in collab. with Mikkelsen Lab, Broad Institute).

- **Population-level and single-cell RNA-seq.** I have applied and developed new computational analyses for RNA-seq expression data from diverse cell types to inform our understanding cell-type heterogeneity and dynamics as well as TF-induced regulation during differentiation. Computational methods for differential analysis have included isoform-prediction and gene count-based analyses as well as graphical assembly-based comparisons for species without a reference genome (e.g. Axolotl). To overcome the difficulty of profiling some cell types due to rarity or the lack of surface markers, a new method was devised for permeabilizing, staining, and FACS sorting cells in a manner compatible with RNA-seq profiling. More recently, I have studied heterogeneity at the single-cell level utilizing the Fluidigm C1 system in pancreatic islets, motor neurons, and assessing dynamics during a prolonged ES differentiation (in my labs and in collab. with Regev Lab, Broad Institute, and Wichterle Lab, Columbia University). To interpret these data new computational methods were developed for quality control, analysis, and interpretation of expression programs. Looking apart from standard RNA-seq preps, which can ignore shorter fragments, an assay was developed to identify lineage-specific short protein coding mRNA.

2003–2011

MIT, Research Assistant, *Bonnie Berger Group & Srinivas Devadas Group (CSAIL)*

Whitehead Institute for Biomedical Research, Research Assistant, *Susan Lindquist Lab.*

- **Amyloid Modeling.** Proteins of all kinds have the ability to adopt amyloid structure: an aggregate state which is often difficult to characterize experimentally, and associated with both disease and beneficial biological function. To study this, I developed new algorithms that predict amyloid structure from sequence highly accurately ($\sim 2\times$ existing methods) by computing an ensemble set of putative conformation variants, modeling *in vivo* structure variability. Further generalized technique to predict energetic landscape differences caused by sequence mutation, and allow simultaneous alignment and prediction of consensus structures given two sequences. Specifically modeled yeast Sup35 proteins to help guide wet experimentation probing structure, and investigated structural and sequence compositional properties of a genomic screen for prion proteins. Commenced inter-institute collaboration with McGill University.
- **Inhibition of biofilms.** Created computational structure models of the functional amyloid Curli in *Enterobacteriaceae*, a protein of unknown structure necessary for biofilm formation. Predicted protein mutations which disrupt biofilm formation (confirmed experimentally by R. Krishnan) and putative inhibitory peptide fragments. Collaboration with R. Krishnan and T. K. Lu experimentally validated mutants and devised a novel phage-display system which inhibits *E. Coli* biofilm formation, colony growth, and invasion of mammalian cells.
- **Transmembrane β -barrel (TMB) Modeling.** Found in outer membranes of bacteria, mitochondria, and chloroplasts, TMB proteins are diversely involved in cell metabolism, but difficult to characterize structurally. For these, I developed algorithms for highly-accurate prediction of TMB structures based on similar ensemble techniques, and a technique for simultaneous alignment and structural prediction of two TMB sequences. Expanded system to include other classes of permuted β -sheet structures. Released web-based software tool (partiFold-TMB/Align), used by dozens of labs for experimentation and bacterial genome annotation.
- **RNA/Protein structure energy models.** Studied methods for constructing pseudo-energy cost functions with minimal parameters. Applied constraint-based linear programming and SVM techniques to reduce the energy model for the RNAfold structure predictor from $\sim 24k$ parameters to ~ 500 , while maintaining similar-to-improved accuracy.

- **SAT-based optimization for bioinformatics.** Designed classical structure prediction and alignment algorithm reductions to SAT instantiations to leverage the increasing power of SMT decision problem solvers.
- **HMM-based secondary structure prediction.** Developed HMM-based protein secondary structure predictor utilizing an intuitive biophysically-motivated pseudo-energy model. Predicted α -helix locations with accuracy unmatched by other tool with as few parameters.
- **Computer architectures for secure processing.** Co-developed AEGIS custom FPGA processor, enabling secure code execution even when processor is physically controlled by an adversary. Required the development of physically unclonable functions (PUFs) and memory integrity verification algorithms. Designed models for partitioning application code into trusted and untrusted components while maintaining secrecy. Introduced methods for secure virtual storage and execution using industry-standard TPM devices.

2006

IBM Research, Summer Internship, *Biomolecular Dynamics and Scalable Modeling Group*

- **BlueMatter molecular dynamics simulator.** Applied concepts from relational databases to massively-parallel scientific MD software, implementing a translation layer that effects data structures and data movements within 16k node BlueGene/L supercomputer. Developed relational primitives described by C annotations providing physical representation flexibility.

2002–2003

Columbia University, Research Assistant, *Steven Nowick Group, Async. Circuits and Systems*

- **Asynchronous boolean logic theory.** Solved open question concerning the existence of hazard-free logic circuits given an arbitrary Boolean function; proving a minimal set of requirements. Designed “extended burst-mode” logic verifier and logical hazard simulator, integrated into CAD tool MINIMALIST.

2003

Columbia University, Research Assistant, *Ken Shepard Group, Integrated Systems Lab.*

- **Signal processing.** Designed digital core and adders for an asynchronous continuous-time sampling digital FIR filter ASIC.

2001–2002

Columbia University, Research Assistant, *Salvatore Stolfo Group, Intrusion Detection Systems*

- **Statistical data-mining.** Devised real-time OS information monitoring techniques for intrusion detection systems and sparse sequence models.

2001

AT&T Research, Summer Internship, *Intelligent Network Services Solutions Group*

- **Decentralized network platform for VoIP.** Developed pluggable modules for computational control of voice calls and Voice-XML dialogues on “ECLIPSE” platform.

2000–2002

Lincoln Electric, Internship, *Machine Research and Development*

- **Embedded weld-precision software.** Assisted design of simulation algorithms and embedded communication software for precision welding.

Publications & Presentations

Journals and Peer-Reviewed Conferences

- L.M.R. Ferreira, T.B. Meissner, T.S. Mikkelsen, W. Mallard, **C.W. O'Donnell**, T. Tilburgs, H.A.B. Gomes, R. Camahort, R.I. Sherwood, D.K. Gifford, J.L. Rinn, C.A. Cowan, J.L. Strominger, A distant trophoblast-specific enhancer controls HLA-G expression at the maternal-fetal interface. Proc. Natl. Acad. Sci., 113(15):5364-5369, 2016.
- S. Hrvatin, **C.W. O'Donnell**, F. Deng, J.R. Millman, F.W. Pagliuca, P. Dilorio, A. Rezanian, D.K. Gifford, D.A. Melton, Differentiated human stem cells resemble fetal, not adult β cells. Proc. Natl. Acad. Sci., 111(8):3038-3043, 2014.
- S. Hrvatin, F. Deng, **C.W. O'Donnell**, D.K. Gifford, D.A. Melton, MARIS: Method for Analyzing RNA following Intracellular Sorting. PLoS ONE, 9(3):e89459, 2014.
- R.I. Sherwood*, T. Hashimoto*, **C.W. O'Donnell***, S. Lewis, A.A. Barkal, J.P. van Hoff, V. Karun, T. Jaakkola, D.K. Gifford (***equal contrib.**), Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape, Nature Biotechnology, 32:171-178, 2014.
- J. Waldispühl*, **C.W. O'Donnell***, S. Will*, S. Devadas, R. Backofen, B. Berger (***equal contrib.**), Simultaneous Alignment and Folding of Multiple Protein Sequences. Journal of Computational Biology, Accepted.
- A. Levin, M. Lis, Y. Ponty, **C.W. O'Donnell**, S. Devadas, B. Berger, J. Waldispühl, A global sampling approach to designing and reengineering RNA secondary structures. Nucleic Acids Research, 40(20):10041-10052, 2012.
- V. Ganesh, **C.W. O'Donnell**, A. Solar-Lezama, S. Devadas, M. Soos, M. Rinard Lynx: A Programmatic SAT Solver for the RNA-folding Problem, 15th International Conference on Theory and Applications of Satisfiability Testing (SAT 2012), Trento, Italy. Lecture Notes in Computer Science 7317, pp. 143-156.
- A.W. Bryan Jr., **C.W. O'Donnell**, M. Menke, L.J. Cowen, S. Lindquist, B. Berger STITCHER: Dynamic assembly of likely amyloid and prion β -structures from secondary structure predictions, Proteins: Structure, Function, and Bioinformatics, 80(2):410-420, 2012.
- R. Halfmann, S. Alberti, R. Krishnan, N. Lyle, **C.W. O'Donnell**, O.D. King, B. Berger, R. Pappu, S. Lindquist, Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins, Molecular Cell, 41(1):72-84, 2011.
- **C.W. O'Donnell**, J. Waldispühl, M. Lis, R. Halfmann, S. Devadas, S. Lindquist, B. Berger, A method for probing the mutational landscape of amyloid structure. 19th Annual International Conference on Intelligent Systems for Molecular Biology & 10th European Conference on Computational Biology (ISMB/ECCB 2011), Vienna, Austria. Bioinformatics, 27(13):i34-42, 2011.
- S. Shenker, **C.W. O'Donnell**, S. Devadas, B. Berger., J. Waldispühl, Efficient traversal of protein folding pathways using ensemble models. Proc. of the 15th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2011), Vancouver, BC, Canada.

- J. Waldispühl, **C.W. O'Donnell**, S. Will, S. Devadas, R. Backofen, B. Berger, Simultaneous Alignment and Folding of Protein Sequences, Proc. of the 13th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2009), Tucson, AZ.
- J. Waldispühl*, **C.W. O'Donnell***, S. Devadas, P. Clote, B. Berger (***equal contrib.**), Modeling Ensembles of Transmembrane β -barrel Proteins, Proteins: Structure, Function, and Bioinformatics, 71(3):1097–1112, 2008.
- B. Gassend, **C.W. O'Donnell**, W. Thies, A. Lee, M. van Dijk, S. Devadas, Learning Biophysically-Motivated Parameters for Alpha Helix Prediction, BMC Bioinformatics, 8(Suppl. 5):S3, 2007.
- B. Gassend, **C.W. O'Donnell**, W. Thies, A. Lee, M. van Dijk, S. Devadas, Predicting Secondary Structure of All-Helical Proteins Using Hidden Markov Support Vector Machines, Proc. of the 2006 Workshop on Pattern Recognition in Bioinformatics, Volume 4146 of LNCS pp. 93-104 (PRIB 2006), Hong Kong, China.
- G.E. Suh, **C.W. O'Donnell**, S. Devadas, AEGIS: A Single-Chip Secure Processor, IEEE Design & Test, 24(6):570–580, 2007.
- **C.W. O'Donnell**, G.E. Suh, M. van Dijk, S. Devadas, Memoization Attacks and Copy Protection in Partitioned Applications, Proc. of the 2007 IEEE Workshop on Information Assurance (IAW 2007), West Point, NY.
- L.F.G. Sarmenta, M. van Dijk, **C.W. O'Donnell**, J. Rhodes, S. Devadas, Virtual Monotonic Counters and Count-Limited Objects using a TPM without a Trusted OS, Proc. of 2006 Workshop on Scalable Trusted Computing (STC 2006), Fairfax, VA.
- G.E. Suh, **C.W. O'Donnell**, S. Devadas, Design and Implementation of the AEGIS Single-Chip Secure Processor Using Physical Random Functions, Proc. of the 32nd Annual International Symposium on Computer Architecture (ISCA 2005), Madison, WI.
- **C.W. O'Donnell**, V. Vaikuntanathan, Information Leak in the Chord Lookup Protocol, Proc. of the 4th Annual IEEE International Conference on Peer-to-Peer Computing (P2P 2004), Zurich.
- S.M. Nowick, **C.W. O'Donnell**, On the Existence of Hazard-Free Multi-Level Logic, Proc. of the 9th Annual IEEE International Conference on Advanced Research in Asynchronous Circuits and Systems (ASYNC 2003), Vancouver.

Other Reports

- **C.W. O'Donnell**, Modeling Ensembles of β -sheet Proteins, MIT EECS Ph.D. Thesis (May 2011).
- **C.W. O'Donnell**, Secure Application Partitioning for Intellectual Property Protection, MIT EECS S.M. Thesis (August 2005).
- G.E. Suh, **C.W. O'Donnell**, S. Devadas AEGIS: A Single-Chip Secure Processor, Elsevier Information Security Technical Report, 10(2):63-73, 2005.
- M. van Dijk, L.F.G. Sarmenta, **C. W. O'Donnell**, S. Devadas Proof of Freshness: How to efficiently use an online single secure clock to secure untrusted memory, MIT CSAIL CSG Technical Memo 481, 2004.

- **C.W. O'Donnell**, G. E. Suh, S. Devadas PUF-Based Random Number Generation, MIT CSAIL CSG Technical Memo 481, 2004.

Patents

- T.K. Lu, R. Krishnan, **C.W. O'Donnell**, S. Devadas, B. Berger, J. Collins, S. Lindquist, Bacteriophages Expressing Amyloid Peptides and Uses Thereof, Provisional applications 61/229,703 (filed July 29, 2009), 61/233,697 (filed August 13, 2009). International Serial Number: PCT/US2010/043770. Published February 3, 2011.

Software

- **PIQ**: Simultaneously identifies genome-wide binding sites for multiple transcription factors based on DNase-seq data. Source/Information: <http://piq.csail.mit.edu/>.
- **AmyloidMutants**: Sequence-based ensemble structure prediction and mutational analysis of amyloid proteins. Webserver: <http://amyloid.csail.mit.edu/>
- **partiFold-TMB**: Sequence-based ensemble structure prediction of transmembrane proteins. Webserver: <http://partifold.csail.mit.edu/TMB/>
- **partiFold-Align**: Simultaneous alignment and structure prediction of transmembrane proteins. Webserver: <http://align.csail.mit.edu/TMB/>
- **MINIMALIST**: Contributed to CAD package for the synthesis and optimization of asynchronous burst-mode controllers.

Teaching Experience

- **Guest Lecturer**, Massachusetts Institute of Technology, Dept. of EECS (Spring 2014)
6.874: Computational Functional Genomics (graduate level): Multiple lectures on the latest research in methods for computational genomic analysis.
- **Guest Lecturer**, Tufts University, Dept. of Computer Science (December 2010)
COMP 150-PBT: Protein Bioinformatic Topics: Guest lecture for Prof. Lenore Cowen on the experimental and computational analysis of amyloid and prion proteins.
- **Co-Instructor**, Massachusetts Institute of Technology, Dept. of Mathematics (Spring 2009)
18.417: Introduction to Computational Molecular Biology (graduate level): Taught introductory algorithms for bioinformatic analysis of proteins and RNAs. Mentored individual students on topics and techniques throughout semester-long research projects.
- **Teaching Assistant**, Massachusetts Institute of Technology, Dept. of Mathematics (Spring 2009)
18.418: Topics in Computational Molecular Biology (graduate level): Managed and advised students on end-of-term research presentations. Presented seminar on protein structure prediction.

- **Teaching Assistant**, Massachusetts Institute of Technology, Dept. of EECS (Fall 2005)
6.170: Software Engineering (undergraduate level): Led weekly group recitations, conducted weekly one-on-one student meetings, assisted full-class lecturing, designed and graded problem sets, mentored 4-person student groups in semester-long research projects, and maintained course infrastructure for homework and project source code.
- **Teaching Assistant**, Columbia University, Dept. of CS (Spring 2002)
CS3824: Computer Organization (undergraduate level): Helped manage and grade course homework material. Held semi-weekly office hours for general computer science assistance.

Awards and Honors

- **Sigma Xi Scientific Research Society**, 2006
- **Theodore R. Bashkow Award**, Columbia University, Computer Science, 2003
Awarded to senior who has excelled in independent research
- **Prentis Scholarship**, Columbia Engineering School Alumni Association, 2003
- **John E. Bertram Memorial Honorary Scholarship**, Columbia University, 2003
- **Dean's List**, Columbia University, 1999-2003
- **Scholarship Winner**, Slovene National Benefits Society, 1999-2003
- **Tau Beta Pi Engineering Honor Society**, 2002
- **Golden Key International Honour Society**, 2002
- **Ralph Haines Honorary Scholarship**, Columbia University, 2001
- **A. Peers Montgomery Memorial Honorary Scholarship**, Columbia University, 2000,2001
- **Scholarship Winner**, American Mutual Life Association, 1999
- **Eagle Scout**, Boy Scouts of America, 1998